

2020-09

Plug-and-play supervisory control using muscle and brain signals for real-time gesture and er...

This work was made openly accessible by BU Faculty. Please [share](#) how this access benefits you. Your story matters.

Version	Published version
Citation (published version):	Joseph DelPreto, Andres F Salazar-Gomez, Stephanie Gil, Ramin Hasani, Frank H Guenther, Daniela Rus. 2020. "Plug-and-play supervisory control using muscle and brain signals for real-time gesture and error detection." Autonomous Robots, Volume 44, Issue 7, pp. 1303 - 1322. https://doi.org/10.1007/s10514-020-09916-x

<https://hdl.handle.net/2144/42733>

Boston University



Plug-and-play supervisory control using muscle and brain signals for real-time gesture and error detection

Joseph DelPreto¹ · Andres F. Salazar-Gomez^{1,2} · Stephanie Gil^{1,3} · Ramin Hasani^{1,4} · Frank H. Guenther⁵ · Daniela Rus¹

Received: 5 December 2018 / Accepted: 9 June 2020 / Published online: 9 August 2020
© The Author(s) 2020

Abstract

Effective human supervision of robots can be key for ensuring correct robot operation in a variety of potentially safety-critical scenarios. This paper takes a step towards fast and reliable human intervention in supervisory control tasks by combining two streams of human biosignals: muscle and brain activity acquired via EMG and EEG, respectively. It presents continuous classification of left and right hand-gestures using muscle signals, time-locked classification of error-related potentials using brain signals (unconsciously produced when observing an error), and a framework that combines these pipelines to detect and correct robot mistakes during multiple-choice tasks. The resulting hybrid system is evaluated in a “plug-and-play” fashion with 7 untrained subjects supervising an autonomous robot performing a target selection task. Offline analysis further explores the EMG classification performance, and investigates methods to select subsets of training data that may facilitate generalizable plug-and-play classifiers.

Keywords Human–robot interaction · EMG control · EEG control · Hybrid control · Gesture detection · Error-related potentials · Plug-and-play supervisory control

1 Introduction

As robots become more prevalent in homes, factories, and other safety-critical settings, detecting and correcting robot errors becomes increasingly important. A fast, reliable, and intuitive framework for supervising robots could help avoid

errors that would otherwise lead to costly hardware damage or safety risks. If a robot could be taught to detect nonverbal cues such as distress signals and hand gestures as reliably as a collaborating human partner, then interactions with robots would become more efficient and supervision or collaboration would become more effective.

Using biosignals such as muscle or brain activity via electromyography (EMG) or electroencephalography (EEG),

This is one of the several papers published in Autonomous Robots comprising the Special Issue on Robotics: Science and Systems.

This work was funded in part by the Boeing Company, for which the authors express gratitude.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10514-020-09916-x>) contains supplementary material, which is available to authorized users.

✉ Joseph DelPreto
delpreto@csail.mit.edu
<https://www.josephdelpreto.com>

Andres F. Salazar-Gomez
salacho@mit.edu

Stephanie Gil
sgil@seas.harvard.edu

Ramin Hasani
ramin.hasani@tuwien.ac.at

Frank H. Guenther
guenther@bu.edu

Daniela Rus
rus@csail.mit.edu

¹ Massachusetts Institute of Technology, Distributed Robotics Lab, Cambridge, MA 02139, USA

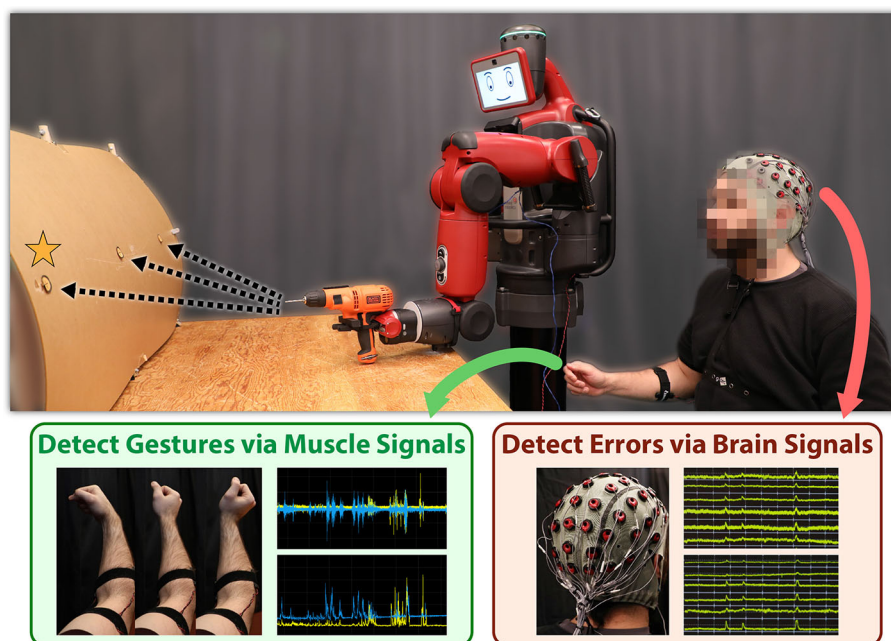
² Present Address: Massachusetts Institute of Technology, Open Learning, Cambridge, MA 02139, USA

³ Present Address: Harvard University, REACT Lab, Cambridge, MA 02138, USA

⁴ Present Address: Technische Universität Wien, Cyber-Physical Systems Group, Vienna, Austria

⁵ Boston University, Guenther, Boston, MA 02215, USA

Fig. 1 A user supervises and controls an autonomous robot, using brain signals to detect mistakes and using muscle signals to correct mistakes by selecting targets. Videos are available at <http://people.csail.mit.edu/delpreto/auro2020>



respectively, has become a promising technique for fast and natural human–robot interaction (HRI). EMG interfaces can control dynamical systems such as exoskeletons, while EEG data can reveal higher-level cognitive states using signals such as the Error-Related Potential (ErrP) indicating a perceived error. Yet reliable real-time detection is challenging since both biosignals are noisy, often difficult to interpret, and vary over time and across subjects. This often leads to the practice of per-subject training phases to tune classifiers, which precludes a “plug-and-play” system where new users can begin controlling the robot from their first interaction.

This paper explores the combination of EEG and EMG into a hybrid interface for robot control. This framework aims to leverage the benefits of each modality to enhance supervisory tasks, especially as classification accuracy continues to improve. The ErrP brain signal is generated unconsciously when a person perceives a mistake and does not need to be taught, facilitating fast passive error detection relying only on user attention. Hand gestures generate characteristic muscle signals that are easier to detect than brain signals, facilitating a vocabulary for reliably indicating desired behavior.

Combining EEG and EMG systems in this way can capitalize on the human’s cognitive ability to judge whether a robot made a mistake and physical ability to indicate correct actions if a mistake is made. As shown in Fig. 1, this paper implements a hybrid control framework in a supervisory scenario where a robot conducts a target selection task for a mock drilling operation. A human supervisor observes the autonomous robot and mentally evaluates whether it chose the correct target. If an ErrP or gesture is detected, the robot halts and requests assistance. The human then gestures left

or right to scroll through possible targets. Once the correct target is selected, the robot resumes autonomous operation.

Two independent classification pipelines process EEG and EMG signals in real time. The EEG pipeline evaluates two neural networks on a buffer of EEG signals, determining whether an ErrP is detected at the moment when the robot begins moving towards a chosen target. The EMG pipeline classifies two channels of surface EMG signals from the forearm on a rolling basis, evaluating a neural network 80 times per second to detect left or right hand-gestures; this allows the user to exert control over the robot at any time and to select a desired target. Both the EEG and EMG classifiers are evaluated in a plug-and-play fashion, only training on data from previous users rather than requiring additional data collection phases for each new supervisor.

Paper contributions This paper works towards using biosignals for effective and reliable supervisory hybrid control in HRI tasks. In particular, its contributions are as follows:

- A framework for combining unconscious error detection via EEG with active error correction via EMG for supervision of autonomous robots during target selection tasks;
- A signal-processing and classification pipeline for continuously detecting left or right hand-gestures based on two surface EMG signals from the forearm, without requiring training on current user data;
- A classification pipeline for detecting unconscious ErrP signals in time-locked EEG signals, evaluated without retraining on each new user;

- *Experimental results from 7 untrained subjects* using the system to supervise a mock drilling operation where the robot chooses from 3 possible targets;
- *Offline EMG analysis* exploring inter-subject muscle signal variations and the impact on classification performance, leading to *a clustering-based algorithm for identifying key subsets of past training data* that could facilitate plug-and-play classifiers for new subjects.

2 Related work

This paper builds on bodies of work investigating human–robot interaction and classification of human biosignals.

2.1 EEG-based methods for human–robot interaction

Many brain-computer interfaces have made progress towards using EEG for communication with both healthy and disabled individuals (Brumberg et al. 2010; Artemiadis and Kyriakopoulos 2011; Birbaumer et al. 1999; Blumberg et al. 2007; Higgin et al. 2015). Using EEG for robotics also shows great promise, with explorations including augmentation of capabilities (Schröer et al. 2015; Penaloza and Nishio 2018), shared control for robotic assistants or grasping (Akinola et al. 2017; Ying et al. 2018), and remote control (Tonin et al. 2011; LaFleur et al. 2013). Yet there are significant challenges such as low signal-to-noise ratios, user training, subject-specific variations, and tuning of detection pipelines (Wolpaw et al. 2002; Vidaurre et al. 2011; McFarland et al. 1997; Lotte and Guan 2011). These issues often lead to added cognitive burden on the user, repeated prompts, or user-specific classification algorithms.

ErrPs are a promising communication mechanism since they occur naturally in response to a perceived error without requiring training or active thought modulation by the human operator (Falkenstein et al. 1991; Schalk et al. 2000; Iturrate et al. 2009, 2010, 2015). ErrPs have been used in brain-computer interfaces and HRI tasks for binary and multi-target classification, correcting classification errors, or controlling robots (Spüler et al. 2012; Wolpaw et al. 1998; Salazar-Gomez et al. 2017; Iturrate et al. 2015; Butfield et al. 2006; Llera et al. 2011; Iturrate et al. 2010; Perrin et al. 2010; Ramli et al. 2015; Zhang et al. 2015; Spüler and Niethammer 2015). Yet these studies also highlight challenges related to real-time classification that precipitate carefully controlled application scenarios and sophisticated detection algorithms (Rivet et al. 2009; Barachant and Bonnet 2011; Barachant et al. 2013; Behncke et al. 2018). Combining ErrP detection with other input signals in a hybrid system could help address some of these challenges while leveraging unconscious error detection.

2.2 EMG-based methods for human–robot interaction

Surface EMG can measure muscle activity via electrodes placed on the skin. Models can then be developed that facilitate controllers based on these signals (Zajac 1989; Hogan 1984; Manal and Buchanan 2003; Ramos and Meggiolaro 2014; Cavallaro et al. 2005; Qashqai et al. 2015; Menon et al. 2016). For example, upper-limb exoskeletons can leverage parameterized muscle models (Ramos and Meggiolaro 2014), neuro-fuzzy systems (Kiguchi and Hayashi 2012), or impedance controllers (Gopura et al. 2009). Such studies have shown that EMG can yield effective human–robot interfaces, but also demonstrate associated challenges including noise, variance between users, and complex muscle dynamics. Some approaches to addressing such challenges include redundant switching models (López et al. 2009) or leveraging the human within the control loop during physical interaction (Lenzi et al. 2011; DelPreto and Rus 2019; Peternel et al. 2017).

Prominent applications of EMG include assistive robots (Gopura et al. 2013) such as exoskeletons (Gopura et al. 2009; Lenzi et al. 2011; Ramos and Meggiolaro 2014; Yin et al. 2012; Kiguchi and Hayashi 2012; Ao et al. 2017; DiCicco et al. 2004; Mulas et al. 2005) and prostheses (Chu et al. 2007; Shenoy et al. 2008). In addition to direct collaboration, muscle signals can be used for remote control or supervision via continuous trajectory control (Artemiadis and Kyriakopoulos 2010; López et al. 2009; Artemiadis and Kyriakopoulos 2011; Artemiadis et al. 2010), augmented reality interfaces (Weisz et al. 2017), or gestures (Crawford et al. 2005; Kim et al. 2008). The presented framework uses a gesture-based EMG control system that allows the human to actively indicate desired robot targets.

2.3 Hybrid control methods for human–robot interaction

Using multiple biosignal sensors can yield hybrid systems that leverage their respective strengths (Müller-Putz et al. 2011). Fusing multiple modalities as well as multiple types of brain signals including ErrPs, researchers have demonstrated promising success in applications for healthy and impaired individuals ranging from quadrotor flight to prostheses (Müller-Putz et al. 2015; Sarasola-Sanz et al. 2017; Kim et al. 2014; Kawase et al. 2017; Ma et al. 2015).

This paper focuses on a framework for hybrid EEG and EMG supervision, which may be applicable to safety-critical tasks where robot operation must be corrected with low latency. Such applications require jointly addressing several challenges including rolling EMG classification to allow human control at arbitrary times, fast online ErrP classifica-

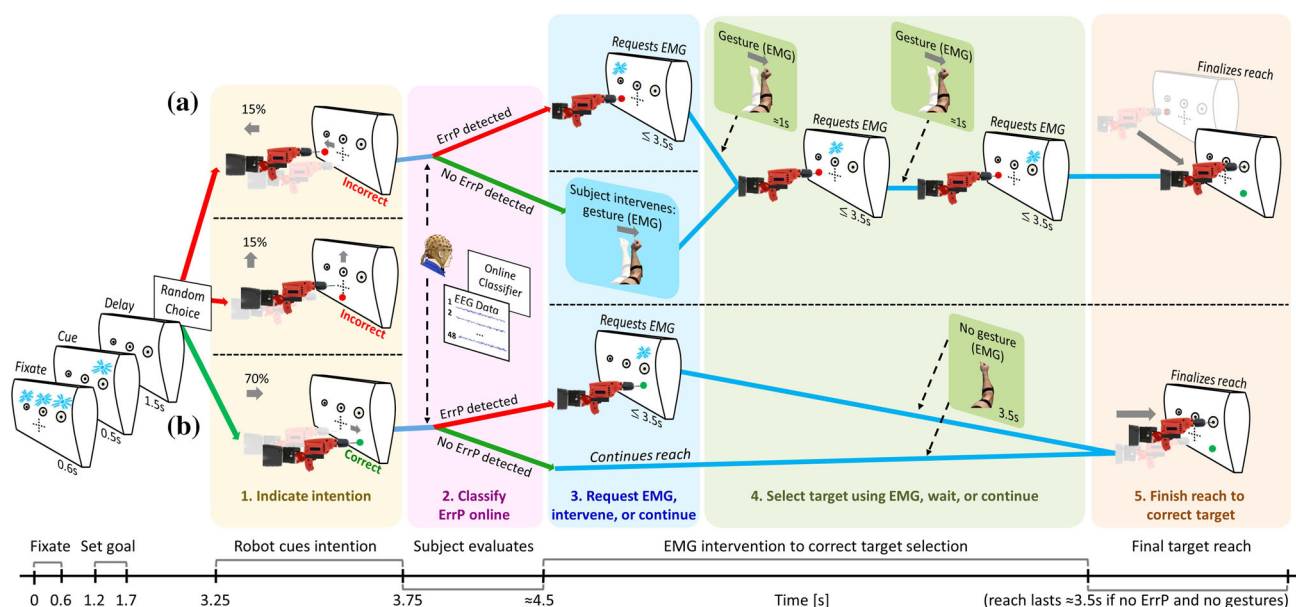


Fig. 2 The main EEG+EMG paradigm is illustrated for when the robot incorrectly chooses the leftmost target (**a**), and when the robot correctly chooses the rightmost target (**b**). The sequence of events when it incorrectly chooses the center target is similar to (**a**). All LEDs blink initially to notify the user of the trial beginning, then one is blinked to indicate the desired drilling location. The robot then moves to indicate its intended target, which is randomly chosen with a bias towards being correct

(block 1); the user mentally evaluates this choice while ErrP classification is performed (block 2). If an ErrP is classified, the robot stops and waits for gestures; otherwise, it continues reaching while searching for EMG intervention (block 3). If a correction is required, the user scrolls through targets via gestures while the robot is stopped (block 4). Upon target selection or no intervention, the robot completes the reach (block 5)

tion even with low signal-to-noise ratios, and experimentally validated system performance in plug-and-play settings.

3 Experimental design

An experiment was designed that allows users to supervise and control an autonomous robot solely via brain and muscle activity. A supervised robot moves an unplugged power drill to one of three targets on a mock plane fuselage, emulating a factory setting where robots assist humans in construction tasks by drilling holes or inserting fasteners.

Two experimental paradigms were implemented: the primary closed-loop supervisory control task, and an open-loop session of gestures. No prior experience controlling a robot or using EEG and EMG is required, encouraging novice users to immediately interact with the robot in more intuitive ways than button sequences or programming.

3.1 Main paradigm: EEG and EMG, closed-loop control

The main task constitutes an autonomous robot selecting targets for a mock drilling operation while a human supervises and intervenes when necessary. As shown in Fig. 1, the human wears an EEG cap and EMG electrodes while sitting

behind the robot and observing the task. The human mentally evaluates whether the robot chooses the correct target in each trial, and uses left or right hand-gestures to correct the robot when necessary. An EEG classifier detects ErrP signals upon initial robot motion, and an EMG classifier continuously identifies gestures.

Figure 2 and the online videos¹ illustrate the sequence of events. Three LEDs are mounted on the mock plane fuselage under the left, center, and right targets. At the start of each trial, the system randomly chooses which target is the desired drilling location with uniform probability. It then illuminates the corresponding LED for 0.5 s to inform the user of the correct location for the upcoming trial. Approximately 1.5 s later the robot randomly chooses a target with a 70% chance of choosing the correct one, and makes an initial arm movement towards the chosen location to cue its intention. The user mentally evaluates whether this motion indicates the correct target, and the EEG classifier assesses whether their brain activity presents an ErrP signal.

If an ErrP is found, the system stops the robot, illuminates the LED representing the robot's currently chosen target, and waits for the human to select the correct target using wrist gestures. A gesture can be a brief flexion (left motion) or extension (right motion) of the right hand. These are used to

¹ <https://people.csail.mit.edu/delpreto/auro2020>.

scroll left or right through the three possible targets; every time a gesture is detected by the EMG classifier, the illuminated LED is changed to indicate the new selection. Since the robot's chosen target can be more than one target away from the correct one, several gestures may be performed. The system considers a target selection finalized 3.5 s after the last detected gesture. All LEDs then turn off, and the robot resumes autonomous operation reaching towards the newly selected target. If an incorrect target was accidentally selected, the supervisor can use gestures to interrupt the robot and initiate another target selection.

If no ErrP signal is found, the robot will continue operating autonomously and reach towards the selected target. However, the user can interrupt the robot by gesturing at any time; this provides a safety feature in case of inaccurate EEG classification. The robot then stops and waits for target selection via EMG as described above. The initial intervention gesture also adjusts the illuminated current selection.

Once the robot reaches the selected target, either with or without intervention, it pauses briefly to indicate completion and then returns to its starting position. This concludes a single trial. On average, an experimental session consisted of 4 blocks of 40 trials each and lasted approximately 2 h.

3.2 EMG-only paradigm: open-loop gestures

Since the main paradigm produces gestures performed at arbitrary times, an EMG-only paradigm was included to generate a corpus of structured EMG training data and facilitate classifier evaluation. However, this data was not used to train a new classifier for the subject that provided it; each experiment used a classifier trained on EMG-only sessions from all previous subjects to implement a plug-and-play strategy. Data from 3 initial subjects that did not participate in online sessions was used as the training set for the first online subject, and was also included in subsequent training sets.

These trials follow a “ready-set-go” sequence to cue time-locked labeled gestures. All three fuselage LEDs blink once to gain the subject's attention, then the left or right LED illuminates for 0.5 s to indicate whether a left or right gesture should be made. After a brief delay, all LEDs illuminate for 1 s; the subject starts and completes their gesture during this period, encouraging consistent gesture timing.

This EMG-only block was performed at the beginning of each experimental session. It consisted of 50 trials and lasted approximately 5 min. The subject is positioned in the same manner as for the main paradigm, but the EEG cap is not worn and the robot is not controlled.

3.3 EEG data paradigm

The EEG classifier was trained on data from 3 sessions of the main closed-loop paradigm performed by 2 preliminary

subjects before the primary set of experiments. No EEG-only paradigm was implemented. Training segments were extracted as the window from 200 ms to 800 ms after the robot begins its initial movement. Subjects were instructed to avoid making gestures during this period to avoid motor-related EEG activity. Although there was a low signal-to-noise ratio during the experiments and a small initial training set, the classifier was kept constant throughout all online sessions to evaluate the hybrid control pipeline in a plug-and-play fashion with a fixed EEG classifier.

3.4 Subject selection

A total of 7 subjects participated in the online control experiments (71.4% male, 85.7% right-handed). In addition, 3 separate subjects provided initial training data; all 3 performed the EMG-only blocks to acquire EMG training data (66.7% male, 66.7% right-handed), and 2 of them performed the main paradigm to acquire EEG training data (50.0% male, 100.0% right-handed). No previous experience using EMG or EEG interfaces was required. Subjects were not screened based on EMG or EEG signals. All subjects provided consent for the study, which was approved by MIT's Committee on the Use of Humans as Experimental Subjects.

4 System overview and data acquisition

An integrated end-to-end system was developed to enable real-time hybrid supervisory control during the target selection tasks. Figure 3 provides an overview of the system.

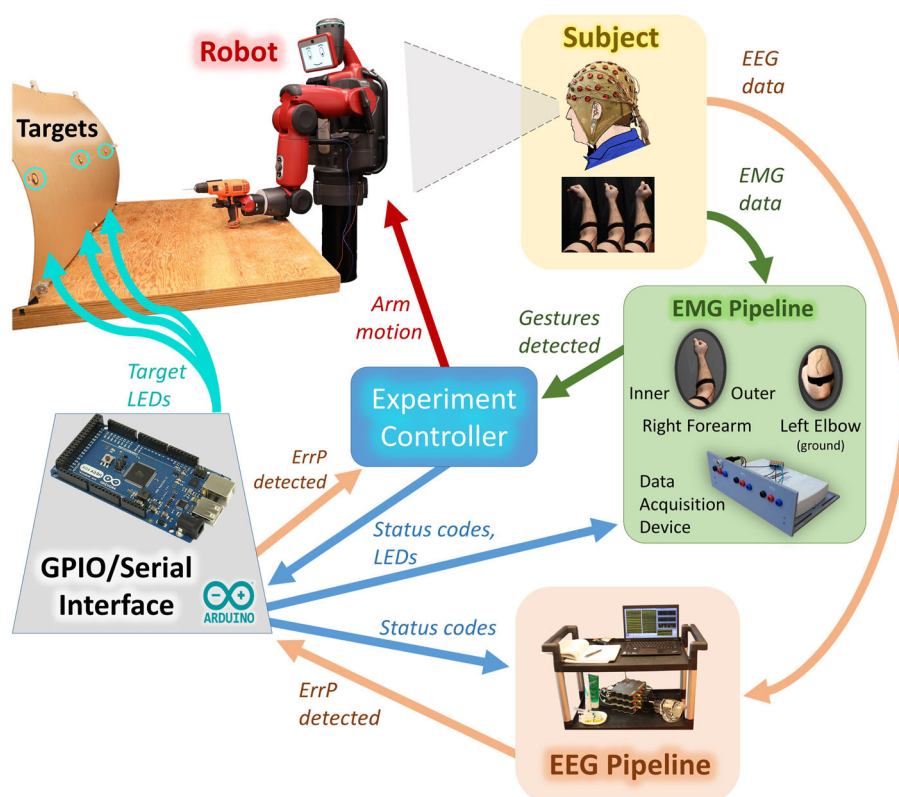
4.1 Experiment controller and robot

The experiment controller, implemented in Python, coordinates all subsystems to realize the paradigms. It chooses correct and selected targets, commands the robot and LEDs, and interprets classifier outputs in the experimental context.

For this particular implementation, the Rethink Robotics Baxter robot was used. It communicates via the Robot Operating System (ROS) with the experiment controller, which provides joint angle trajectories for Baxter's left 7 degree-of-freedom arm. A pushbutton switch is fastened under the arm to determine exactly when the arm lifts from the table; this is used as a time-locking signal for EEG acquisition.

An Arduino Mega 2560 serves as an interface between the experiment controller and the classification subsystems. Key events such as trial timing, chosen targets, robot motion, and LED states are sent to the Arduino via USB serial. These are mapped to predefined 7-bit code words, which the Arduino uses to set a parallel port. The 8th pin of the port is wired to the pushbutton switch on Baxter's arm. The EMG and EEG data acquisition systems read this port along with their

Fig. 3 The system includes EMG and EEG acquisition and classification systems, an experiment controller, and the Baxter robot. A human supervisor closes the loop



respective biosignals, allowing for synchronization between data and experimental events.

4.2 EMG hardware and data acquisition

Two differential pairs of reusable non-adhesive surface bar electrodes are placed over the user's right posterior (outer) forearm and right anterior (inner) forearm, positioned over the muscles using recommendations in De Luca (2002). An additional electrode is placed slightly distal to the left elbow as a ground reference. Electrode placement sites are cleaned with Nuprep skin preparation gel and alcohol prep pads to reduce electrical impedance and improve adhesion. Conductive gel is applied beneath each electrode. An elastic Velcro strap around each forearm holds the electrodes in place.

An NI USB-6216 data acquisition (DAQ) device is connected directly to the electrodes, the Arduino parallel port, and the LED control signals. The 16-bit analog input channels are configured for a -0.2V to $+0.2\text{V}$ range. Differential mode is used to reduce common noise. Analog signals are sampled at 2000 Hz and sent via USB as buffers of 200 samples every 0.1 s. These buffers are acquired by Simulink (2017b), which performs online signal processing and gesture identification. Classifications are sent asynchronously to the experiment controller via ROS.

4.3 EEG hardware and data acquisition

A total of 48 passive electrodes, following the 10-20 scalp distribution, are used for EEG data collection. Three Guger Technologies USBamps sample all signals at 256 Hz. Ground and reference electrodes are placed at the AFz position and the right ear, respectively. The Arduino parallel port is connected directly to GPIO inputs of the USBamps.

Online signal processing and classification is performed in Simulink (2015a). The pushbutton switch on Baxter's arm initiates an EEG buffer, as described in Salazar-Gomez et al. (2017). The buffer is then processed and passed to the ErrP classifier. Classifications are sent asynchronously to the experiment controller via the Arduino, using a USBamp GPIO output.

5 Classification of EMG and EEG signals

Two independent classification pipelines were implemented: one for continuous gesture detection from muscle signals, and one for time-locked error detection from brain signals. Each one operated online and was trained only on data from previous users.

Algorithm 1 EMG Processing, Training, and Classification

SIGNAL PROCESSING

- 1: *raw* \leftarrow sample 2 differential EMG signals at 2 kHz
- 2: *filtered* \leftarrow band-pass filter 5–400 Hz: FIR high-pass (stopband 2 Hz, passband 5 Hz) + FIR low-pass (passband 400 Hz, stopband 450 Hz)
- 3: *amplified* \leftarrow amplify by 1,000
- 4: *envelope* \leftarrow rectify, low-pass (FIR 2nd-order, passband 5 Hz), amplify by 1.5
- 5: *downsampled* \leftarrow downsample each channel to 80 Hz
- 6: **if** preparing training data (EMG-only blocks) **then**
- 7: *centered* \leftarrow extract segment, center around peak
- 8: *scaled* \leftarrow shift each min to 0, jointly normalize
- DATA AUGMENTATION
- 9: **if** gesture segment (one per trial) **then**
- 10: *newPositiveExamples* \leftarrow shift segment slightly left and right
- 11: *newNegativeExamples* \leftarrow shift segment greatly left and right
- 12: **end if**
- 13: **if** baseline segment (two per trial) **then**
- 14: *newNegativeExample* \leftarrow shift segment greatly left or right
- 15: **end if**
- GENERATE FEATURE VECTORS
- 16: *trimmed* \leftarrow truncate segments to 1.2 s
- 17: *featureVectors* \leftarrow concatenate 2 channels (total 192 samples)
- TRAIN SINGLE-LAYER NEURAL NETWORK
- 18: *network* \leftarrow 20-neuron hidden layer, 3 outputs
- 19: **else if** predicting online (EMG-only or EEG+EMG) **then**
- 20: *buffers* \leftarrow rolling 1.2 s window of each channel
- 21: *scaled* \leftarrow shift each min to 0, jointly normalize
- 22: *featureVectors* \leftarrow concatenate 2 channels (total 192 samples)
- PREDICT LEFT/RIGHT/BASELINE LABEL
- 23: *classification* \leftarrow max neural network output
- 24: *prediction* \leftarrow filter network classifications
- 25: **end if**

5.1 EMG classification: continuous gesture detection

Muscle signals are acquired from the inner and outer right forearm as described in Sect. 4.2. These two signals are passed through a pipeline of signal processing, feature extraction, and classification to detect gestures on a rolling basis. This pipeline is outlined in Algorithm 1.

5.1.1 Signal processing

Each muscle signal is independently band-pass filtered, amplified, and envelope-detected. The initial filtering preserves the useful frequency content of the EMG signal (De Luca 2002) while removing DC offsets, low-frequency motion artifacts, and high-frequency noise. The envelopes indicate muscle activation levels, as shown in Fig. 4.

5.1.2 Segmentation and normalization

As described in Sect. 3.2, EMG training data was collected by cueing subjects to make left or right gestures during specified time windows using LEDs. The data was then segmented according to those LED signals. To accommodate variable reaction times and gesture durations, each extracted segment begins 0.75 s before LEDs turn on and ends 0.75 s after LEDs

turn off. This yields one labeled gesture segment per trial. Two baseline segments without gestures were also extracted from each trial when LEDs were off.

For each extracted segment, each EMG channel's envelope is shifted so its minimum value is at 0. Both envelopes are then scaled by the same factor so the peak value becomes 1. Finally, they are downsampled to 80 Hz. The right column of Fig. 4 presents sample results. This shifting and scaling helps standardize inputs across subjects and time, making classification robust to variations in EMG magnitude and offset. Normalizing each segment independently alleviates issues of calibration, fatigue, and gesture variations.

5.1.3 Data augmentation

To detect gestures on a rolling basis, the trained network should be robust to small time shifts while preferring gestures that are centered in the buffer. This helps make predictions smooth and reliable while eliminating false detections. For example, the top row of Fig. 4 illustrates antagonistic muscle activity during a left gesture that may result in predicting two different gestures, a left gesture then a right gesture, if the network is too insensitive to time shifts.

A data augmentation approach, demonstrated in Fig. 5, was used to guide the network towards robust rolling classification. Each extracted gesture segment is first centered around the peak value. Two copies are then synthesized by shifting slightly left and right by random amounts (1–100 ms); these are assigned the original gesture label. Two copies are also synthesized by shifting farther left and right (400 ms plus a random amount up to 100 ms); these are assigned a baseline label. By creating slightly shifted positive examples and greatly shifted negative examples, this augmentation guides the network towards preferring gestures that are centered in the buffer within a specified tolerance.

For each original baseline segment, a single synthetic baseline example is extracted via a far shift.

Each example is then truncated to 1.2 s (96 samples), using a window centered around the original centered segment. If there is not enough data in the original segment to perform one of the shifts, that synthetic example is discarded.

5.1.4 Neural network training

As a result of the segment extraction and data augmentation, each trial yields 9 training examples: 3 positive gesture examples, 2 baseline examples with shifted gestures, and 4 baseline examples without gestures. The training corpus is thus biased towards negative examples, which is acceptable since avoiding false gesture detections is important for smooth operation and the rolling online classifier will encounter vastly more baseline segments than gestures.

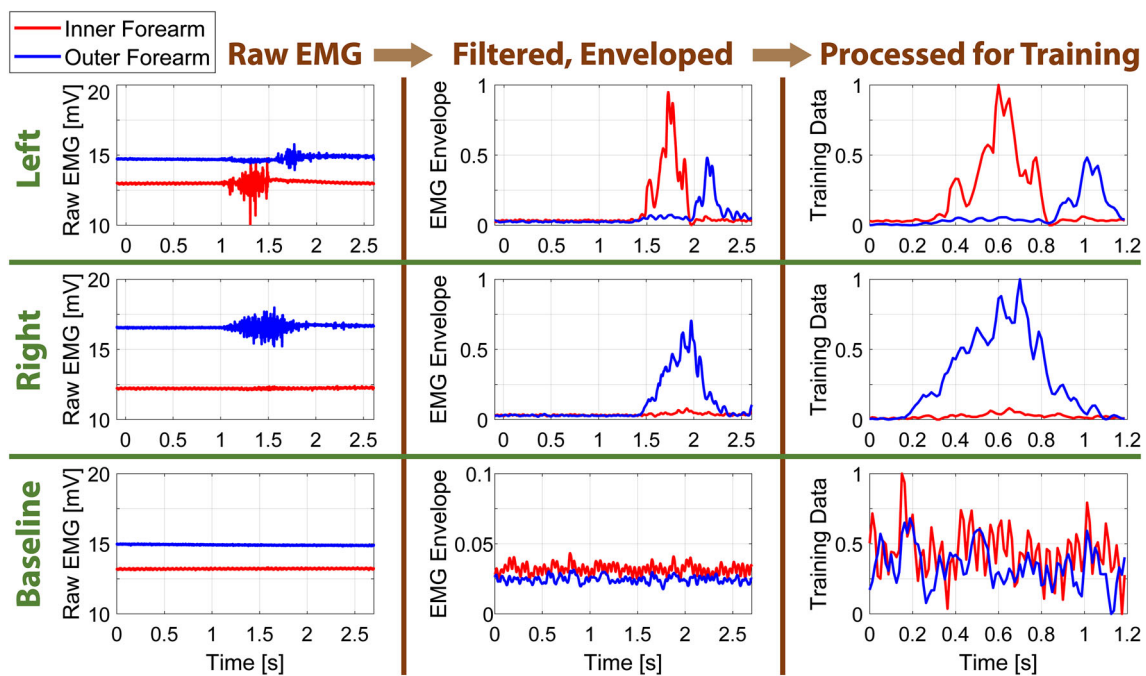


Fig. 4 Acquired EMG signals are passed through a signal-processing pipeline. Raw signals for randomly selected left-gesture, right-gesture, and baseline segments are shown on the left. Detected envelopes are

shown in the center. The right column shows the segments after shifting down to 0, normalizing, trimming, centering, and downsampling

The two EMG envelopes within each example are concatenated to yield a feature vector with 192 elements. These labeled vectors are used to train a feed-forward neural network, using the Pattern Recognition functionality of Matlab's Neural Network Toolbox (2017b). The network has a single hidden layer of size 20 using a hyperbolic tangent sigmoid activation function, and an output layer of size 3 using a softmax activation function. The 3 outputs are used to classify a segment as baseline, a left gesture, or a right gesture.

A new classifier was trained for each experiment using data from the EMG-only blocks of all previous subjects, including the 3 preliminary subjects, without using data from the current subject.

5.1.5 Online continuous classification

As the two muscle signals are acquired by Simulink, the signal-processing pipeline is applied and the detected envelopes are downsampled to 80 Hz. These downsampled envelopes populate rolling buffers of duration 1.2 s. Each time the buffers are updated, they are independently shifted down to 0, jointly normalized, and concatenated. The resulting 192-sample vector is passed to the trained neural network. To avoid spurious predictions, network classifications are slightly filtered. A rolling buffer of 12 network classifications (150 ms) is maintained, and a gesture is declared if at least 60% of them are not baseline and at least 60% of them

are the same label. Rising edges in the stream of filtered classifications are interpreted as new gesture predictions.

5.2 EEG classification

Brain signals are acquired as described in Sect. 4.3, then passed through a pipeline of signal processing, feature extraction, and classification to detect ErrP signals.

5.2.1 EEG pre-processing and feature extraction

A buffer of EEG data is initiated during each closed-loop trial when the robot begins its arm motion (at stimulus onset), as detected by a pushbutton switch under the robot's arm. This buffer collects 800 ms of data from the 48 EEG channels. A decoding window from 200 ms to 800 ms post stimulus onset is then extracted. These signals are band-pass filtered for 1–10 Hz using a 4th-order zero-phase Butterworth filter. Based on offline analysis, only signals from 9 electrodes on the mid-line central region, corresponding to the locations FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2, were selected. These 9 filtered channels are concatenated to create a 1,386-element feature vector for classification.

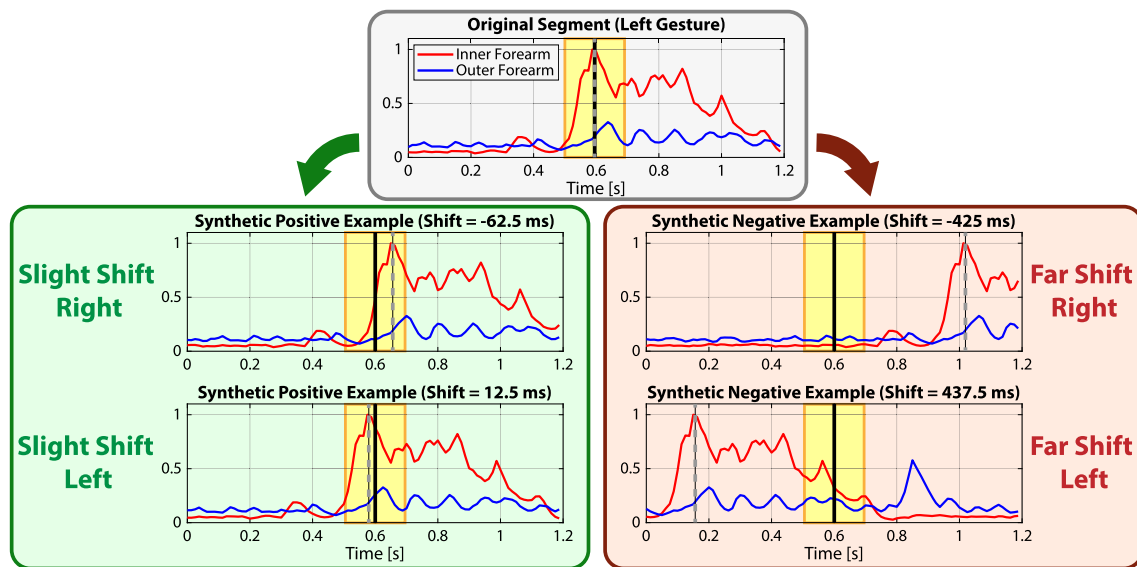


Fig. 5 Training data is augmented via time-shifting, to encourage the network to prefer centered gestures within a specified tolerance (shaded yellow regions). Each gesture trial is shifted slightly to create new positive examples, and shifted farther to create new “baseline” examples

5.2.2 Network training and online ErrP classification

The EEG classification pipeline was trained on 3 preliminary sessions from 2 of the preliminary subjects. The classifier then remained constant for all 7 online sessions to evaluate the hybrid control pipeline in a plug-and-play fashion.

Two feed-forward neural networks were trained and evaluated. The first network has a single hidden layer (input-100-1). The second network has four hidden layers (input-100-50-100-10-1). To perform binary ErrP classification, a threshold was chosen for each network by minimizing the following cost function:

$$Cost = \sqrt{(1 - sensitivity)^2 + (1 - specificity)^2} \quad (1)$$

The offline area under the curve (AUC) metrics for the simpler and deeper neural networks were 70% and 69%, respectively. Offline analysis averaging the regression outputs from both networks and using an averaged threshold increased performance by 3 percentage points. Thus, the final classification pipeline implemented in Simulink uses both networks: the feature vector is fed in parallel to each classifier, then the two outputs are averaged and compared to the averaged threshold. This final output is sent to the experiment controller as the ErrP detection flag.

6 Experimental results and discussion

The system was used for 7 sessions, each with a different subject. This enabled evaluation of the interface efficacy as well as the EMG and EEG classifiers.

6.1 System performance: integrated hybrid control

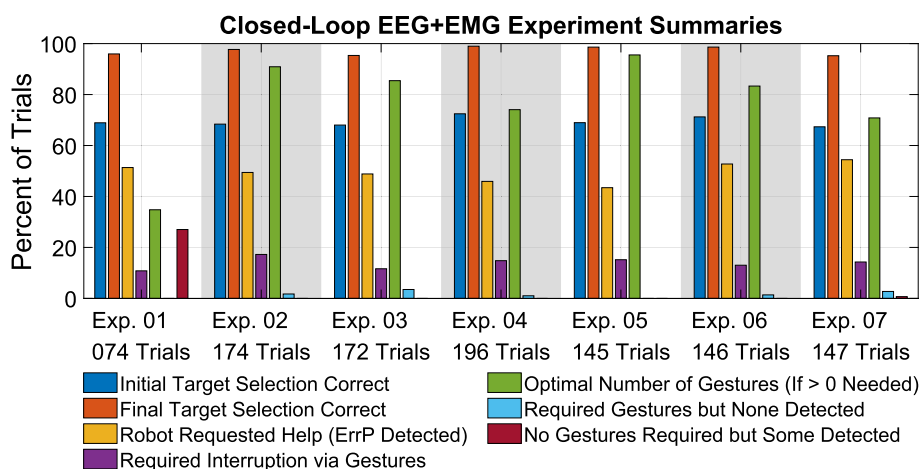
Results summarizing the trials and exploring effectiveness of the overall system are shown in Fig. 6. There were 151 trials per experiment on average, after removing trials in which subjects self-reported being distracted. The robot randomly chose the correct target in 69.5% of the trials, and after EEG and EMG control chose the correct target in 97.3% of the trials. The hybrid control thus allowed the user to correct errors made by the robot and by the EEG classifier. In most interventions, the minimum number of gestures required to scroll to the correct target was detected by the system even though subjects were not instructed to minimize their gestures; this suggests efficient usage of the gesture interface and a lack of spurious false gesture classifications.

Compared to fully autonomous trials with neither error nor gestures detected, trials in which an error was detected and the user performed gestures to select a new target averaged 8.2 s longer with standard deviation (SD) 2.8 s. Trials in which no error was detected via EEG but the user interrupted the robot via gestures averaged 5.6 s (SD 2.4 s) longer than fully autonomous trials.

6.2 EMG classification performance

The EMG classification pipeline was evaluated during the EMG-only blocks and during the main-paradigm blocks. Online results are presented below, while Sect. 7 presents collected data and offline analysis.

Fig. 6 Overall system performance is summarized by whether the robot placed the drill at the correct target after hybrid EEG and EMG control. In addition, the user's interaction and some of the failure modes are described by considering when gestures were performed or required



6.2.1 Open-loop gesture detection: EMG-only blocks

Figure 7 summarizes the rolling classification performance during open-loop gesture blocks. As described in Sect. 3.2, each trial includes a single cued gesture. Classifiers trained on previous blocks were run online as described in Sect. 5.1.5, but no feedback was presented to the user. The classifiers made a single correct gesture prediction in 92.8% of the 345 trials. There were mixed correct and incorrect gesture predictions in 3.2% of trials, no non-baseline predictions in 3.2% of trials, and multiple correct-gesture predictions in 0.9% of trials. There were no trials in which a left or right gesture was classified as the opposite gesture without additional correct predictions. These results indicate that the classifiers robustly and accurately detected gestures during EMG-only blocks.

As described in Sect. 5.1.5, the neural network classifications are slightly filtered. Without this filter, the pipeline would have made a single correct gesture prediction in 77.7% of trials and multiple correct-gesture predictions in 15.4% of trials. This indicates that the filter aided performance by decreasing repeated correct predictions, but was not needed to remove incorrect predictions.

6.2.2 Closed-loop gesture detection: EMG+EEG blocks

The classifiers trained on past EMG-only blocks were also used during the closed-loop trials in which users could arbitrarily make left or right gestures at any time. Ground truth gesture labels were obtained by annotating videos of the subjects' arms in post-processing. Video was not recorded during the first EMG+EEG experiment, so closed-loop EMG performance results are not available for that experiment.

Figure 8 depicts per-subject and aggregated confusion matrices. The classifiers correctly identified 65.8% of left gestures and 85.2% of right gestures, while not falsely identifying any right gestures as left gestures and only 1.8% of left gestures as right gestures. Most instances of missing a

gesture occurred when the subject made multiple rapid gestures, such that there were likely multiple gestures within the rolling buffer window.

The bottom row of Fig. 8a evaluates prediction of baseline activity. The experiments spanned over 18,000 s of predictions being generated at 80 Hz, and there were 17 gesture predictions when no gesture was performed. This spans the entire experiment, including between trials when subjects could move and reposition themselves. It is thus promising that gestures were very rarely predicted when no gesture was intended by the subject.

On average, left-gesture motions lasted for 0.85 s (SD 0.26 s) and were detected 1.15 s (SD 0.16 s) after motion initiation. Right-gesture motions lasted for 0.79 s (SD 0.23 s) and were detected 1.09 s (SD 0.11 s) after motion initiation. Thus, the classifiers generally waited to predict a gesture until it was contained by and relatively centered in the 1.2 s buffer window. Once a detection commenced, a left or right prediction was continuously outputted for an average of 0.20 s or 0.26 s (both SD 0.08 s), respectively, demonstrating smooth predictions. Together, these results imply that the data augmentation successfully encouraged the networks to prefer centered gestures within a specified tolerance – to be robust to small but not large time shifts.

Overall, these results indicate that the EMG classifiers provided a reliable plug-and-play method of communicating with the robot via gestures. The low false positive rate indicates that the robot rarely stopped unnecessarily. If a gesture was missed or misidentified, the subject could simply make another gesture to correct the target selection. The detection latency is also reasonable for real-time control and closed-loop feedback. These results demonstrate effective supervisory control in a multiple-choice context, and are promising for future extensions of the gesture vocabulary beyond the two well-defined hand motions currently explored.

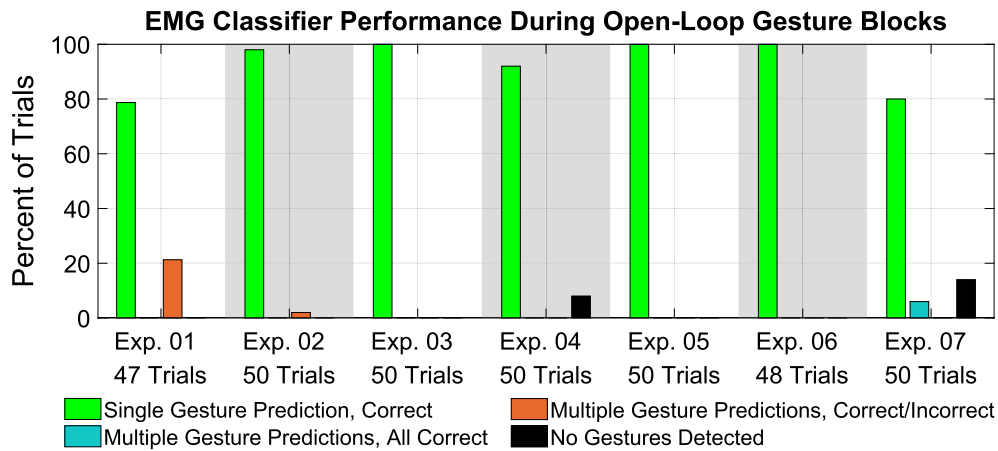
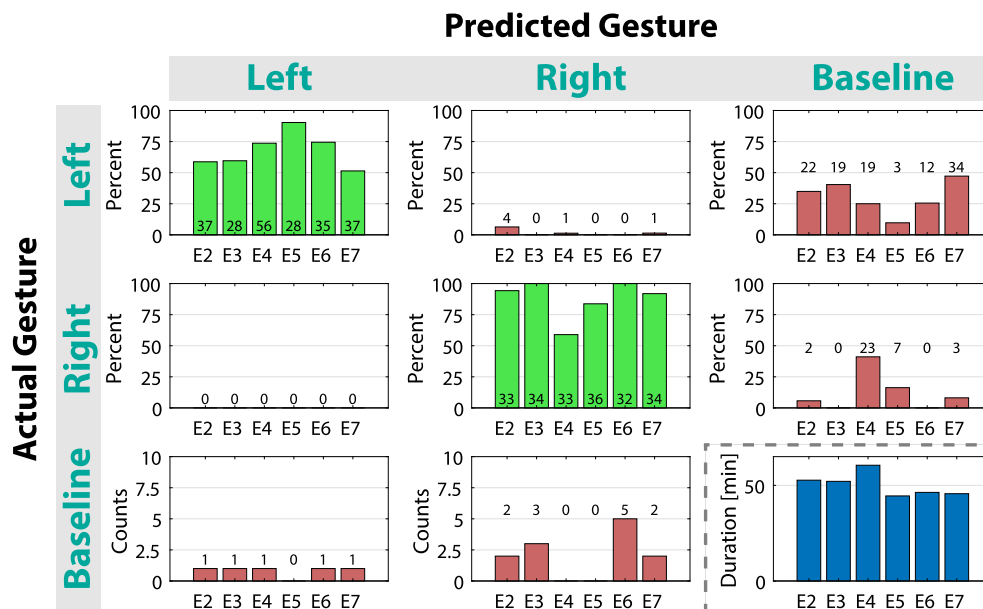


Fig. 7 In each experiment, an EMG classifier trained on previous subjects was continuously invoked during the open-loop EMG-only trials. The results indicate successful real-time gesture detection and generalization to new subjects



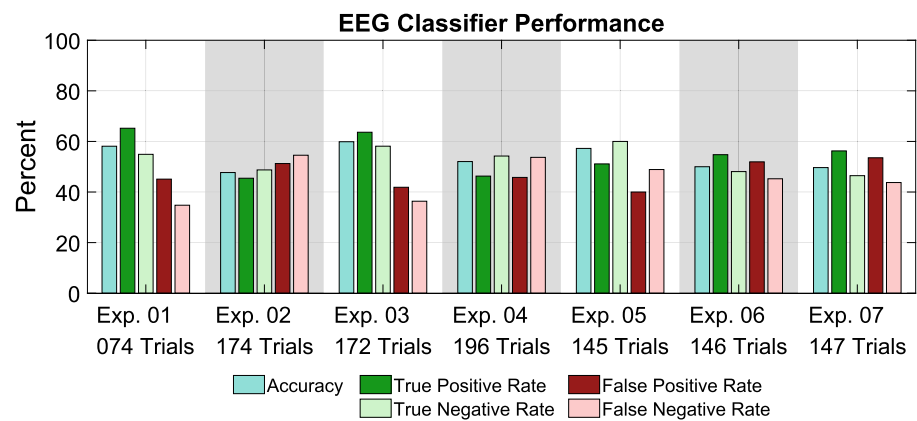
(a) Each entry of the confusion matrix is decomposed to illustrate per-experiment classification performance. The number on each bar represents the number of trials associated with that percent.

		Predicted Gesture					
		Left		Right		Baseline	
Actual Gesture		Count	Percent	Count	Percent	Count	Percent
	Left	221	65.8%	6	1.8%	109	32.4%
	Right	0	0.0%	202	85.2%	35	14.8%
	Baseline	5		12		Total Experimental Duration:	
Total		226		220		301.7 min	

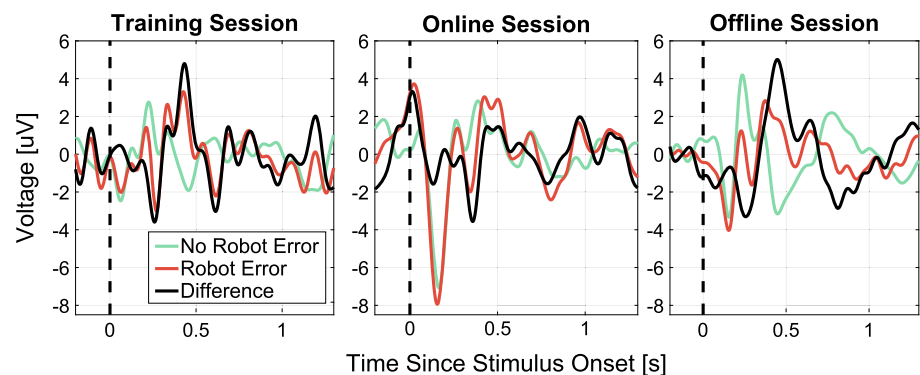
(b) A single confusion matrix for all subjects summarizes performance.

Fig. 8 Confusion matrices summarize the EMG classification performance during closed-loop EMG+EEG blocks. Users could make left or right gestures at any time

Fig. 9 EEG results include online classification performance and a visual comparison of the ErrP signals collected during training, testing, and previously recorded sessions



(a) The EEG classification performance is summarized by the overall accuracy, true positive rate, and true negative rate in each experiment.



(b) EEG traces representing the average of the trials from a representative subject's data (electrode channel Cz) are presented from the preliminary training set (left), the online experiments used for testing (center), and a previous set of experiments (Salazar-Gomez et al. 2017) (right). Averages are shown for trials that contained a robot error (red), that did not contain a robot error (green), and for the difference between them (black).

6.3 EEG classification performance

Figure 9a summarizes the online EEG classification performance in each experiment. Although it was lower than the EMG classification performance, it was sufficient for preliminary investigation of the presented framework for combining EMG and EEG into a hybrid control interface. Using the more reliable EMG interface to correct errors by the EEG interface was acceptable in the current task, but future studies with more time-critical tasks should improve EEG performance to better leverage unconscious error detection. In particular, the current EEG system used a classifier that was trained on only 3 sessions from 2 subjects and that was not updated after each experiment; a larger corpus of ErrP data would likely yield a more robust classifier, as suggested by results of previous work (Salazar-Gomez et al. 2017).

Improving the EEG signal quality may also help improve classification accuracy. Figure 9b presents averaged traces from three different scenarios: the training set, the online experiments, and a related robot supervision task explored in previous work (Salazar-Gomez et al. 2017) (referred to as the offline session). The known negative-positive-negative structure of an ErrP that is expected in the difference traces (black traces) is seen in the offline sessions, but it is less apparent in the training sessions or especially in the online sessions. The online sessions generally presented lower signal-to-noise ratios than expected. This suggests that reducing the noise of the raw EEG signal during online sessions might yield improved classification results for future experiments. Many EEG studies carefully separate subjects from sources of interference such as robots and actuators, but the current experiments featured subjects quite close to the robot since the focus was on testing the capabilities of the hybrid control

system; the impact of this placement and other signal quality concerns could be further evaluated in future investigations.

7 Offline EMG analysis for plug-and-play training

The amount of training data needed as well as how to choose the most informative data among past subjects are crucial questions for creating a plug-and-play gesture prediction system. After completing the online experiments, collected EMG data for left and right gestures was used to explore these questions and further evaluate the system's generalizability in plug-and-play evaluations.

Section 7.1 outlines the training and testing procedures used during all of the presented offline analyses. Section 7.2 then investigates gesture variations between individual subjects and the viability of classifiers trained on a single subject, while Sect. 7.3 investigates training on multiple subjects and how the size of the training group affects classification performance. Building on these results, Sect. 7.4 presents a clustering-based algorithm for strategically selecting training data from past subjects that may facilitate creating plug-and-play classifiers for future subjects.

7.1 Offline training and evaluation procedure

Data from all of the EMG-only open-loop gesture blocks was used. The 7 subjects that performed online experiments will be referred to as subjects 1–7, and the 3 preliminary subjects will be referred to as subjects 8–10. Mirroring the procedure used for online experiments, the offline analyses train networks on time-locked gesture examples but evaluate using rolling classification.

Training examples were extracted, processed, and augmented as described in Sect. 5.1. For each individual classifier used in offline analysis, 10 neural networks with the same architecture as used during online experiments were trained. Each one randomly divided gesture examples from the subset of subjects being considered by the analysis into training, validation, and testing examples using a 70–15–15 split. The one with the highest accuracy on its testing set was then selected for use in the analysis. This offline training process mimics what could be done in an online experimental context when creating a neural network for use with a future unknown subject.

Offline performance scores were then obtained by simulating the selected networks in a streaming fashion. The complete online pipeline described in Sect. 5.1.5 was executed with recorded raw EMG data and the new classifiers. Since the data is from open-loop sessions, this process mimics what would have been observed online. Accuracy is then reported as the fraction of trials in which these new

rolling classifications indicate a single gesture and indicate the correct gesture (predicting multiple gestures in a trial is considered a failure even if they all indicate the cued gesture).

7.2 Inter-subject variability

Examining variability between subjects' gestures and the generalizability of training on individual subjects can yield insight into creating plug-and-play classifiers. Figure 10 reveals commonalities among all subjects such as the general signal shape and the primary muscles. Yet it also reveals distinctive traits such as the prominence of antagonistic muscle activity, gesture speed, and consistency across trials.

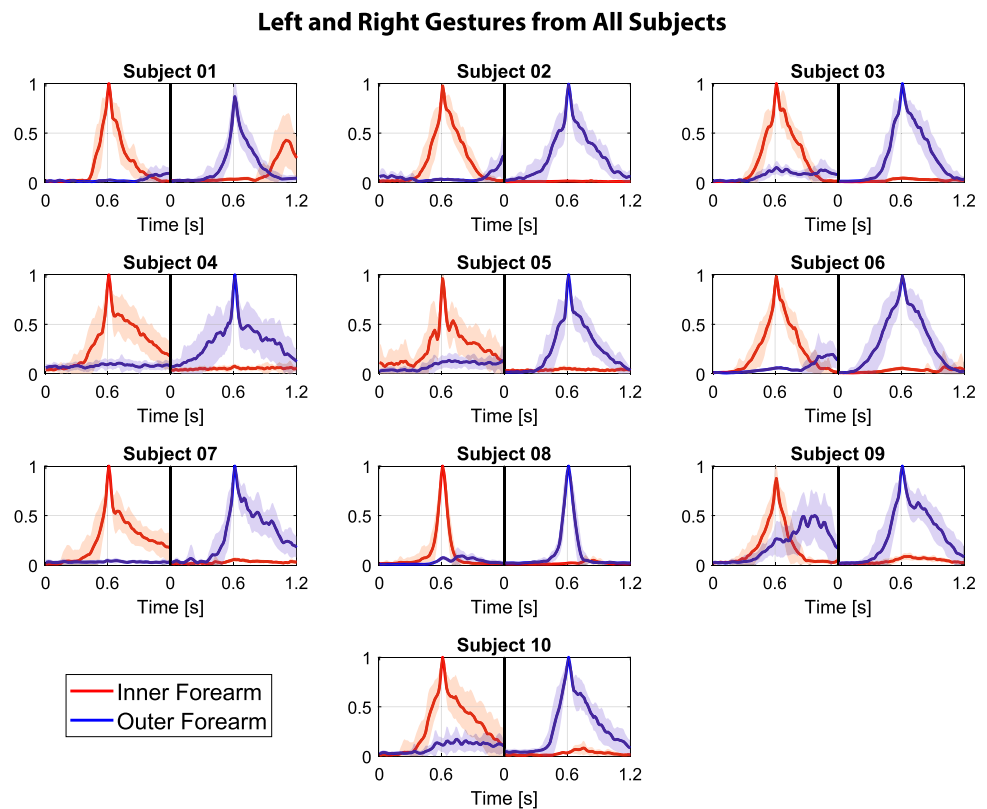
To investigate the impact of these variations on system performance, classifiers were trained on data from a single subject and then tested on each of the remaining 9 subjects. Figure 11a aggregates results by the exclusive training subject used. It illustrates that training on certain subjects can be more generalizable than others; for example, training on subjects 1 or 2 yields higher accuracies than training on subjects 5, 7, 8, 9, or 10 ($p < 0.05$ for each of these 10 pairwise comparisons). In addition to trends regarding training on specific subjects, Fig. 11b reveals trends regarding testing on specific subjects. For example, testing on subjects 3, 5, 6, or 10 yields higher accuracies across all training subjects than testing on subjects 1, 2, 8, or 9 ($p < 0.04$ for each of these 16 pairwise comparisons). The performance matrix is also not symmetric. These results demonstrate that it can be difficult to predict how a subject's data will impact the neural network and its performance on future subjects. Yet they also suggest that a small training set may be sufficient to create a generalizable classifier if chosen carefully.

7.3 Number of training subjects

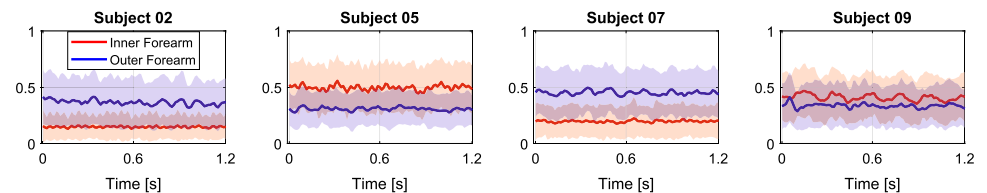
Building on results from individual subjects, performance and robustness were explored when combining data from multiple subjects. A neural network was trained on every possible grouping of subjects, using groups of size 1 through 9. Each network was evaluated on each subject not in its training group. This yielded $\sum_{N=1}^9 \frac{10!}{N!(10-N)!} = 1022$ neural networks and $\sum_{N=1}^9 \frac{10!(10-N)}{N!(10-N)!} = 5110$ evaluations. This whole process was then repeated 5 times, to help account for randomness in the neural network training process.

Figure 12a summarizes the results. Training on all previous subjects (groups of 9 subjects) yielded the highest mean accuracy of 91.2% (SD 12.4%). Training on a single subject yielded a lower distribution of accuracies than each other group size ($p < 0.01$ for each of the 8 pairwise comparisons). However, pairwise statistical comparisons among all of the accuracy distributions do not reveal a consistent pattern based on the number of training subjects; the current

Fig. 10 Training segments extracted from EMG-only blocks are illustrated for all subjects. Thicker lines represent mean traces, and shading spans 1 standard deviation on each side. The segments have been vertically shifted, normalized, downsampled, and centered around their peak. Synthetic augmentation examples are not included



(a) Processed left and right gestures, separated by black vertical lines, are illustrated for each subject to demonstrate trends and variability.



(b) The baseline segments of 4 selected subjects are illustrated.

results therefore suggest that the amount of training data did not have a reliable impact on performance beyond a single subject, but increasing the sample sizes via additional experiments would be needed to further investigate such trends. The results also suggest that a small number of training subjects might be sufficient if chosen strategically.

7.4 Choosing training subjects via clustering

The previous analyses suggest that training on certain subjects may be more generalizable than others, and that using small groups of subjects may be desirable. However, it is still unclear how to choose training subjects that can leverage the similarities and variations between their gestures to improve

performance on future subjects. Towards this end, a subject selection algorithm based on clustering is explored.

To first evaluate the effectiveness of clustering in the context of the collected muscle signals, clustering was performed on gestures from all 10 subjects. This was done on left and right gestures independently, using *k*-means clustering (Lloyd 1982) with correlation as a distance metric. An optimal number of clusters was chosen using the silhouette evaluation metric of Matlab (2018b) on up to 10 clusters.

As illustrated in Fig. 13, two clusters were found for left gestures and three clusters were found for right gestures. The prominence of antagonistic muscle activity appears to be identified as a distinctive feature for both gestures. The speed of motion onset and return to neutral also seems to be a distinctive feature for right gestures. These traits agree

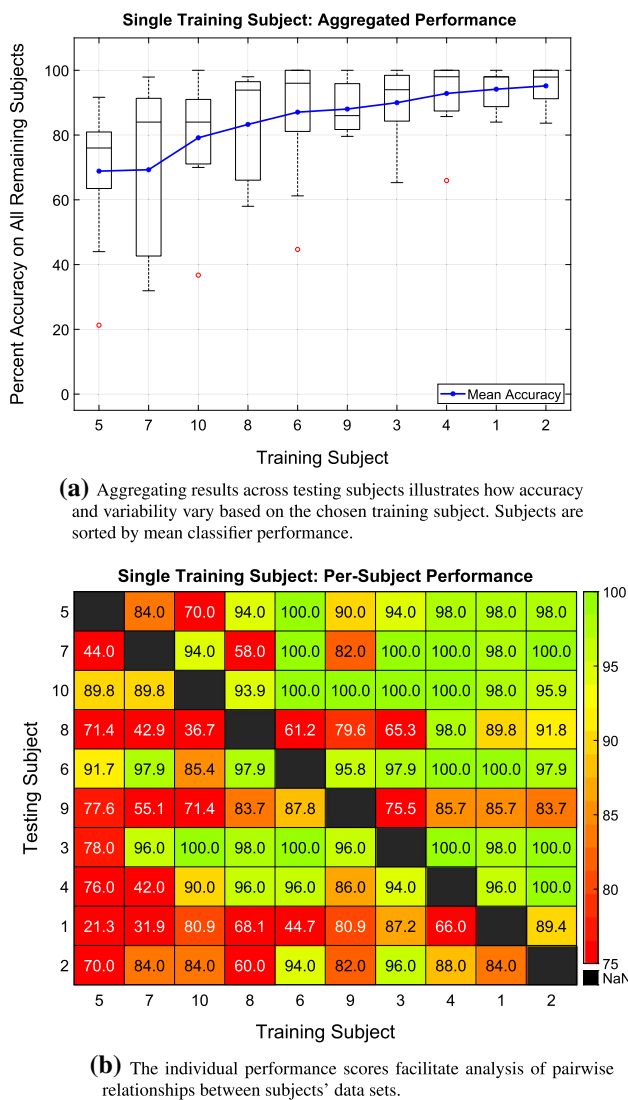


Fig. 11 Training a classifier on each subject individually and then separately evaluating on each remaining subject investigates the generalizability of gestures from each subject. Each box in (a) aggregates the corresponding column of (b)

with the discussion of Fig. 10, and suggest that the chosen clustering mechanism is reasonable for the current task.

A greedy algorithm was then implemented to apply this identification of distinctive gesture characteristics to training subject selection. Left and right gestures from past subjects are first clustered as described above. Training subjects are then greedily selected until each cluster has at least 30% of its gestures included in the training pool; at each iteration, the unselected subject with the most gestures from the least-represented cluster is selected and all gestures from that subject are added to the training pool. This aims to prefer subjects that exemplify certain gestural characteristics well; less consistent subjects with gestures from multiple clusters are less likely to be selected for training. Using a fixed cluster

coverage ratio helps preserve the original gesture breakdown, so more common characteristics are still more prevalent in the downselected corpus.

This algorithm was evaluated with a leave-one-subject-out strategy to simulate plug-and-play performance; each subject was in turn treated as an unknown future subject. For each subject left out, the clustering and greedy subject selection was performed on the 9 known subjects. A neural network trained on the chosen subjects was then evaluated on the unknown subject that was not part of the clustering and selection procedure. This whole process was repeated 5 times for each subject left out, to help account for randomness in the clustering and training processes.

Results are summarized in Fig. 12b. Accuracy averaged 94.4% (SD 9.2%) across all subjects left out. The distribution of accuracies was significantly higher than the first 8 distributions of Fig. 12a ($p < 0.02$ for each of these 8 pairwise comparisons). While increasing the sample sizes would be needed to further investigate such performance comparisons, the current results also suggest that using the clustering-based selection algorithm generally achieved comparable or improved performance on new subjects compared to what would have been achieved by training on all past subjects, despite using less training data.

The algorithm chose 4.2 training subjects on average (SD 0.4) from the 9 available, which agrees with the conclusions drawn from Fig. 12a regarding the feasibility of a small training group size if the subjects can be chosen strategically. The mean optimal number of left and right gesture clusters was 2.2 (SD 0.5) and 2.5 (SD 0.6), respectively, which is consistent with the clustering of Fig. 13.

Subjects 7, 8, and 9 were chosen most frequently, being included in 100.0%, 95.6%, and 80.0% respectively of the iterations in which they were not the subject left out. Subjects 1, 6, 3, and 10 were included in 51.1%, 51.1%, 46.7%, and 15.6% of such iterations, respectively, and the remaining subjects were chosen in less than 15.0% of such iterations.

These results suggest that the presented clustering-based algorithm may facilitate generalizable classifiers for future subjects by identifying a key subset of past subjects on which to train. While the present analysis uses a vocabulary of two well-defined gestures and a relatively small number of subjects, the results are promising for future investigations with more subjects and for extensions to additional scenarios.

8 Conclusion and future work

Detection, correction, and prevention of robot errors are important tasks for a human–robot interface to accomplish. Moving towards these goals, the presented system combines brain and muscle signals into a hybrid framework for detecting and correcting robot mistakes during selection tasks.

Fig. 12 For every possible subset of subjects ranging in size from 1 to 9 subjects, 5 classifiers were trained and then independently evaluated on each of the remaining subjects; results are then grouped by the size of the training subset (a). The cluster-based algorithm for selecting training subjects can yield improved performance on new subjects (b)

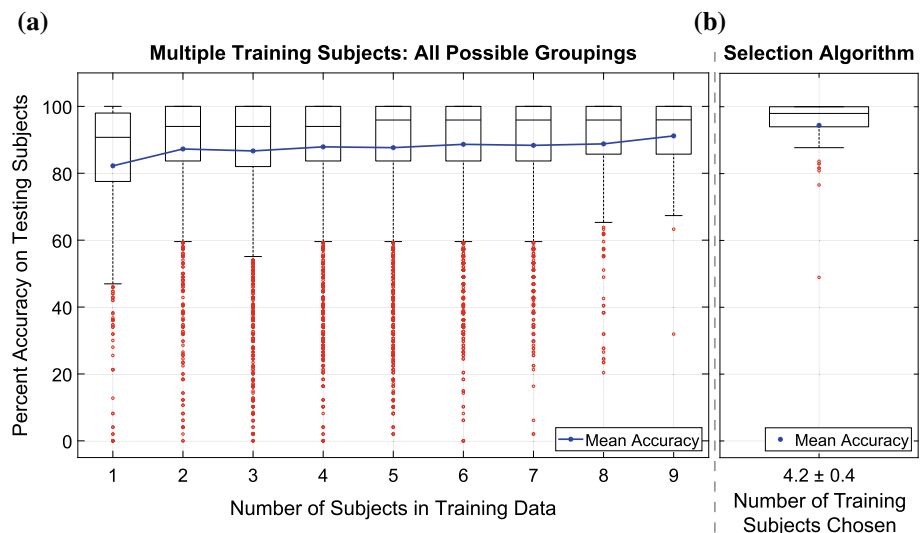
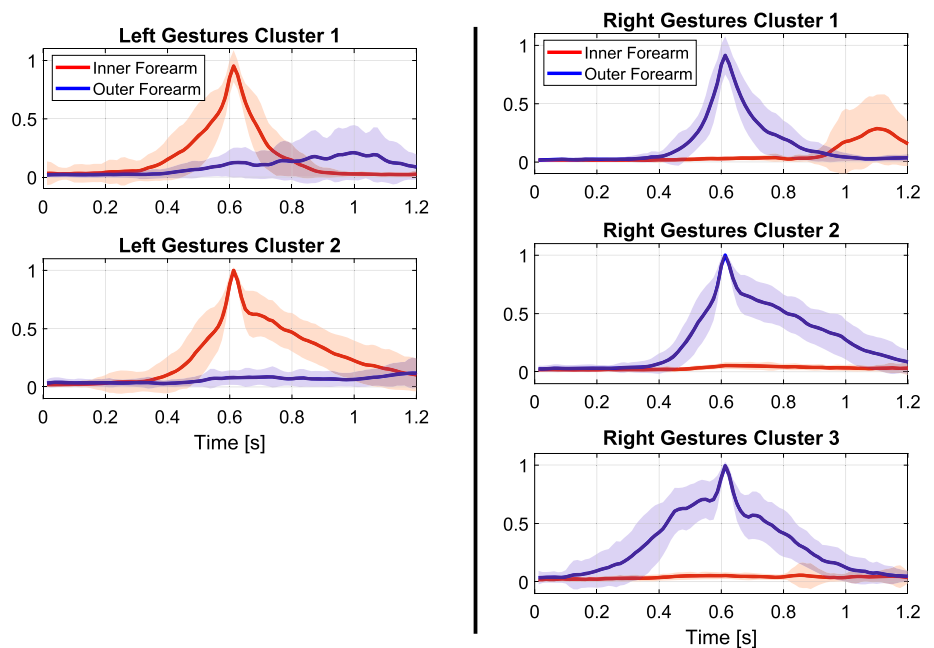


Fig. 13 Performing *k*-means clustering on left and right gestures from all subjects yields two left-gesture clusters and three right-gesture clusters. Bold traces represent cluster centroids, and shading represents 1 standard deviation on each side. The clustering highlights antagonistic muscle activity and gesture onset speed as key characteristics



Error-related potential signals in the brain can provide an unconscious mechanism for quickly detecting when a user perceives a robot error, and detecting gestures via muscle signals can provide a reliable mechanism for actively indicating desired behavior. Both pipelines were evaluated with 7 untrained subjects in a plug-and-play fashion to reduce the barrier to new users controlling the robot.

While this system has demonstrated the use of a hybrid EMG+EEG interface for robot control, future work is required to investigate whether it could be deployed in safety-critical or time-critical tasks. In particular, the EEG classification accuracy should be improved beyond the 54% currently observed in the robot control experiments. Increasing the size of the EEG training corpus to include significantly more than 2 subjects, improving the EEG signal quality, and

potentially using other biosignals such as EMG as a training input for the EEG system may help increase the ErrP classification performance and reliability in future studies.

Future work can also investigate this framework with a larger subject pool and with additional robot control tasks. Evaluating the EMG pipeline on more users can further investigate the observed trends in inter-subject variations and their impact on performance, and can further test the performance of the presented clustering-based subject selection algorithm for training plug-and-play classifiers. The presented selection algorithm could also be adapted to choose individual gesture examples instead of treating subjects' datasets as atomic units. Future work could also increase the gesture vocabulary beyond the two hand gestures currently implemented

to address a broader range of supervisory tasks and test the generalizability of the presented detection approach.

This work thereby moves towards improved human–robot interaction in situations where effective supervision can mean the difference between a dangerous environment and a safe one, or between a costly mistake and a swift intervention.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Artemiadis, P. K., & Kyriakopoulos, K. J. (2011). A switching regime model for the EMG-based control of a robot arm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 41(1), 53–63. <https://doi.org/10.1109/TSMCB.2010.2045120>.
- Akinola, I., Chen, B., Koss, J., Patankar, A., Varley, J., & Allen, P. (2017). Task level hierarchical system for bci-enabled shared autonomy. In *2017 IEEE-RAS 17th international conference on humanoid robotics (Humanoids)*, IEEE, pp 219–225. <https://doi.org/10.1109/HUMANOIDS.2017.8246878>.
- Ao, D., Song, R., & Gao, J. (2017). Movement performance of human-robot cooperation control based on EMG-driven hill-type and proportional models for an ankle power-assist exoskeleton robot. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(8), 1125–1134. <https://doi.org/10.1109/TNSRE.2016.2583464>.
- Artemiadis, P. K., & Kyriakopoulos, K. J. (2010). EMG-based control of a robot arm using low-dimensional embeddings. *IEEE Transactions on Robotics*, 26(2), 393–398. <https://doi.org/10.1109/TRO.2009.2039378>.
- Artemiadis, P. K., & Kyriakopoulos, K. J. (2010). An EMG-based robot control scheme robust to time-varying EMG signal features. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), 582–588. <https://doi.org/10.1109/TITB.2010.2040832>.
- Barachant, A., & Bonnet, S. (2011). Channel selection procedure using Riemannian distance for BCI applications. In *2011 International IEEE/EMBS conference Neural engineering (NER)*, IEEE, pp. 348–351. <https://doi.org/10.1109/NER.2011.5910558>.
- Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2013). Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing*, 112, 172–178. <https://doi.org/10.1016/j.neucom.2012.12.039>.
- Behncke, J., Schirrmester, R. T., Burgard, W., & Ball, T. (2018). The signature of robot action success in eeg signals of a human observer: Decoding and visualization using deep convolutional neural networks. In *2018 6th international conference on brain-computer interface (BCI)*, IEEE, pp. 1–6. <https://doi.org/10.1109/IWW-BCI.2018.8311531>.
- Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., et al. (1999). A spelling device for the paralysed. *Nature*, 398(6725), 297–298. <https://doi.org/10.1038/18581>.
- Blumberg, J., Rickert, J., Waldert, S., Schulze-Bonhage, A., Aertsen, A., & Mehring, C. (2007). Adaptive classification for brain computer interfaces. In *2007 Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 2536–2539. <https://doi.org/10.1109/IEMBS.2007.4352845>.
- Brumberg, J. S., Nieto-Castanon, A., Kennedy, P. R., & Guenther, F. H. (2010). Brain-computer interfaces for speech communication. *Speech Communication*, 52(4), 367–379. <https://doi.org/10.1016/j.specom.2010.01.001>.
- Buttfield, A., Ferrez, P. W., & Millan, J. R. (2006). Towards a robust BCI: error potentials and online learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2), 164–168. <https://doi.org/10.1109/TNSRE.2006.875555>.
- Cavallaro, E., Rosen, J., Perry, J. C., Burns, S., & Hannaford, B. (2005). Hill-based model as a myoprocessor for a neural controlled powered exoskeleton arm - parameters optimization. In *Proceedings of the 2005 IEEE international conference on robotics and automation (ICRA)*, IEEE, pp. 4514–4519. <https://doi.org/10.1109/ROBOT.2005.1570815>.
- Chu, J. U., Moon, I., Lee, Y. J., Kim, S. K., & Mun, M. S. (2007). A supervised feature-projection-based real-time EMG pattern recognition for multifunction myoelectric hand control. *IEEE/ASME Transactions on Mechatronics*, 12(3), 282–290. <https://doi.org/10.1109/TMECH.2007.897262>.
- Crawford, B., Miller, K., Shenoy, P., & Rao, R. (2005). Real-time classification of electromyographic signals for robotic control. In *Proceedings of AAAI (Vol. 5, pp. 523–528)*. <http://aaai.org/Papers/AAAI/2005/AAAI05-082.pdf>.
- De Luca, C. J. (2002). Surface electromyography: Detection and recording. DelSys Incorporated. https://www.delsys.com/Attachments_pdf/WP_SEMGintro.pdf.
- DelPreto, J., & Rus, D. (2019). Sharing the load: Human-robot team lifting using muscle activity. In *IEEE international conference on robotics and automation (ICRA)*. IEEE. <https://doi.org/10.1109/ICRA.2019.8794414>.
- DiCicco, M., Lucas, L., & Matsuoka, Y. (2004). Comparison of control strategies for an EMG controlled orthotic exoskeleton for the hand. In *2004 IEEE international conference on robotics and automation (ICRA)* IEEE, Vol. 2, pp. 1622–1627. <https://doi.org/10.1109/ROBOT.2004.1308056>.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. ii. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology* 78(6), 447–455. [https://doi.org/10.1016/0013-4694\(91\)90062-9](https://doi.org/10.1016/0013-4694(91)90062-9).
- Gopura, R., Bandara, D., Gunasekara, J., & Jayawardane, T. (2013). *Recent Trends in EMG-Based Control Methods for Assistive Robots*. IntechOpen: Electrodiagnosis in new frontiers of clinical research. <https://doi.org/10.5772/56174>.
- Gopura, R. A. R. C., Kiguchi, K., & Li, Y. (2009). SUEFUL-7: a 7DOF upper-limb exoskeleton robot with muscle-model-oriented EMG-based control. In *2009 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, IEEE, pp. 1126–1131. <https://doi.org/10.1109/IROS.2009.5353935>.
- Higger, M., Akcakaya, M., Nezmfar, H., LaMountain, G., Orhan, U., & Erdogmus, D. (2015). A bayesian framework for intent detection and stimulation selection in SSVEP BCIs. *IEEE Signal Processing Letters*, 22(6), 743–747. <https://doi.org/10.1109/LSP.2014.2368952>.
- Hogan, N. (1984). Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Transactions on Automatic Control*, 29(8), 681–690. <https://doi.org/10.1109/TAC.1984.1103644>.

- Iturrate, I., Antelis, J. M., Kubler, A., & Minguez, J. (2009). A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. *IEEE Transactions on Robotics*, 25(3), 614–627. <https://doi.org/10.1109/TRO.2009.2020347>.
- Iturrate, I., Montesano, L., & Minguez, J. (2010). Single trial recognition of error-related potentials during observation of robot operation. In *2010 annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 4181–4184. <https://doi.org/10.1109/IEMBS.2010.5627380>.
- Iturrate, I., Chavarriaga, R., Montesano, L., Minguez, J., & Millán, J. R. (2015). Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control. *Scientific Reports*, 5, 13893. <https://doi.org/10.1038/srep13893>.
- Kawase, T., Sakurada, T., Koike, Y., & Kansaku, K. (2017). A hybrid BMI-based exoskeleton for paretic: EMG control for assisting arm movements. *Journal of Neural Engineering*, 14(1), 016015. <https://doi.org/10.1088/1741-2552/aa525f>.
- Kiguchi, K., & Hayashi, Y. (2012). An EMG-based control for an upper-limb power-assist exoskeleton robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42(4), 1064–1071. <https://doi.org/10.1109/TSMCB.2012.2185843>.
- Kim, B. H., Kim, M., & Jo, S. (2014). Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking. *Computers in Biology and Medicine*, 51, 82–92. <https://doi.org/10.1016/j.combiomed.2014.04.020>.
- Kim, J., Mastnik, S., & André, E. (2008). EMG-based hand gesture recognition for realtime biosignal interfacing. In *Proceedings of the 13th international conference on intelligent user interfaces*, ACM, pp. 30–39. <https://doi.org/10.1145/1378773.1378778>.
- LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., & He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface. *Journal of Neural Engineering*, 10(4), 046003. <https://doi.org/10.1088/1741-2560/10/4/046003>.
- Lenzi, T., De Rossi, S., Vitiello, N., & Carrozza, M. (2011). Proportional EMG control for upper-limb powered exoskeletons. In *2011 Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 628–631. <https://doi.org/10.1109/IEMBS.2011.6090139>.
- Llera, A., van Gerven, M. A., Gómez, V., Jensen, O., & Kappen, H. J. (2011). On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks*, 24(10), 1120–1127. <https://doi.org/10.1016/j.neunet.2011.05.006>.
- Lloyd, S. (1982). Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>.
- López, N. M., di Sciascio, F., Soria, C. M., & Valentinuzzi, M. E. (2009). Robust EMG sensing system based on data fusion for myoelectric control of a robotic arm. *Biomedical Engineering Online*, 8(1), 5. <https://doi.org/10.1186/1475-925X-8-5>.
- Lotte, F., & Guan, C. (2011). Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2), 355–362. <https://doi.org/10.1109/TBME.2010.2082539>.
- Ma, J., Zhang, Y., Cichocki, A., & Matsuno, F. (2015). A novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs: Application to robot control. *IEEE Transactions on Biomedical Engineering*, 62(3), 876–889. <https://doi.org/10.1109/TBME.2014.2369483>.
- Manal, K., & Buchanan, T. S. (2003). A one-parameter neural activation to muscle activation model: estimating isometric joint moments from electromyograms. *Journal of Biomechanics*, 36(8), 1197–1202. [https://doi.org/10.1016/S0021-9290\(03\)00152-0](https://doi.org/10.1016/S0021-9290(03)00152-0).
- McFarland, D. J., McCane, L. M., David, S. V., & Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, 103(3), 386–394. [https://doi.org/10.1016/S0013-4694\(97\)00022-2](https://doi.org/10.1016/S0013-4694(97)00022-2).
- Menon, S., Migimatsu, T., & Khatib, O. (2016). A parameterized family of anatomically accurate human upper-body musculoskeletal models for dynamic simulation & control. In *2016 IEEE-RAS 16th international conference on humanoid robots (Humanoids)*, pp. 587–594. <https://doi.org/10.1109/HUMANOIDS.2016.7803334>.
- Mulas, M., Folgheraiter, M., & Gini, G. (2005). An EMG-controlled exoskeleton for hand rehabilitation. In *2005 International conference on rehabilitation robotics (ICORR)*, IEEE, pp. 371–374. <https://doi.org/10.1109/ICORR.2005.1501122>.
- Müller-Putz, G., Leeb, R., Tangermann, M., Höhne, J., Kübler, A., Cincotti, F., et al. (2015). Towards noninvasive hybrid brain-computer interfaces: framework, practice, clinical application, and beyond. *Proceedings of the IEEE*, 103(6), 926–943. <https://doi.org/10.1109/JPROC.2015.2411333>.
- Müller-Putz, G. R., Breitwieser, C., Cincotti, F., Leeb, R., Schreuder, M., Leotta, F., et al. (2011). Tools for brain-computer interaction: a general concept for a hybrid bci. *Frontiers in Neuroinformatics*, 5, 30. <https://doi.org/10.3389/fninf.2011.00030>.
- Penaloza, C. I., & Nishio, S. (2018). BMI control of a third arm for multitasking. *Science Robotics*, 3(20), eaat1228. <https://doi.org/10.1126/scirobotics.aat1228>.
- Perrin, X., Chavarriaga, R., Colas, F., Siegwart, R., & Millán, J. R. (2010). Brain-coupled interaction for semi-autonomous navigation of an assistive robot. *Robotics and Autonomous Systems*, 58(12), 1246–1255. <https://doi.org/10.1016/j.robot.2010.05.010>.
- Peternel, L., Tsagarakis, N., & Ajoudani, A. (2017). A human-robot co-manipulation approach based on human sensorimotor information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(7), 811–822. <https://doi.org/10.1109/TNSRE.2017.2694553>.
- Qashqai, A., Ehsani, H., & Rostami, M. (2015). A hill-based emg-driven model to estimate elbow torque during flexion and extension. In *2015 22nd Iranian conference on biomedical engineering (ICBME)*, IEEE, pp. 166–171. <https://doi.org/10.1109/ICBME.2015.7404136>.
- Ramli, R., Arof, H., Ibrahim, F., Mokhtar, N., & Idris, M. Y. I. (2015). Using finite state machine and a hybrid of EEG signal and EOG artifacts for an asynchronous wheelchair navigation. *Expert Systems with Applications*, 42(5), 2451–2463. <https://doi.org/10.1016/j.eswa.2014.10.052>.
- Ramos, J. L. A., & Meggiolaro, M. A. (2014). Use of surface electromyography for human amplification using an exoskeleton driven by artificial pneumatic muscles. In *Biomedical robotics and biomechatronics, 2014 IEEE RAS & EMBS international conference*, IEEE, pp. 585–590. <https://doi.org/10.1109/BIOROB.2014.6913841>.
- Rivet, B., Soulloumiac, A., Attina, V., & Gibert, G. (2009). xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8), 2035–2043. <https://doi.org/10.1109/TBME.2009.2012869>.
- Salazar-Gomez, A. F., DelPreto, J., Gil, S., Guenther, F. H., & Rus, D. (2017). Correcting robot mistakes in real time using EEG signals. In *2017 IEEE international conference on robotics and automation (ICRA)*, IEEE, pp. 6570–6577. <https://doi.org/10.1109/ICRA.2017.7989777>.
- Sarasola-Sanz, A., Irastorza-Landa, N., López-Larraz, E., Bibián, C., Helmhold, F., Broetz, D., Birbaumer, N., & Ramos-Murguialday, A. (2017). A hybrid brain-machine interface based on EEG and EMG activity for the motor rehabilitation of stroke patients. In *2017 International conference on rehabilitation robotics (ICORR)*, IEEE, pp. 895–900. <https://doi.org/10.1109/ICORR.2017.8009362>.

- Schalk, G., Wolpaw, J. R., McFarland, D. J., & Pfurtscheller, G. (2000). EEG-based communication: presence of an error potential. *Clinical Neurophysiology*, 111(12), 2138–2144. [https://doi.org/10.1016/S1388-2457\(00\)00457-0](https://doi.org/10.1016/S1388-2457(00)00457-0).
- Schröder, S., Killmann, I., Frank, B., Völker, M., Fiederer, L., Ball, T., & Burgard, W. (2015). An autonomous robotic assistant for drinking. In *2015 IEEE international conference on robotics and automation (ICRA)*, IEEE, pp. 6482–6487. <https://doi.org/10.1109/ICRA.2015.7140110>.
- Shenoy, P., Miller, K. J., Crawford, B., & Rao, R. P. (2008). Online electromyographic control of a robotic prosthesis. *IEEE Transactions on Biomedical Engineering*, 55(3), 1128–1135. <https://doi.org/10.1109/TBME.2007.909536>.
- Spüler, M., & Niethammer, C. (2015). Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity. *Frontiers in Human Neuroscience*, 9, 155. <https://doi.org/10.3389/fnhum.2015.00155>.
- Spüler, M., Bensch, M., Kleih, S., Rosenstiel, W., Bogdan, M., & Kübler, A. (2012). Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI. *Clinical Neurophysiology*, 123(7), 1328–1337. <https://doi.org/10.1016/j.clinph.2011.11.082>.
- Tonin, L., Carlson, T., Leeb, R., & Millán, JdR. (2011). Brain-controlled telepresence robot by motor-disabled people. In *2011 annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 4227–4230. <https://doi.org/10.1109/IEMBS.2011.6091049>.
- Vidaurre, C., Sannelli, C., Müller, K. R., & Blankertz, B. (2011). Co-adaptive calibration to improve BCI efficiency. *Journal of Neural Engineering*, 8(2), 025009. <https://doi.org/10.1088/1741-2560/8/2/025009>.
- Weisz, J., Allen, P. K., Barszap, A. G., & Joshi, S. S. (2017). Assistive grasping with an augmented reality user interface. *The International Journal of Robotics Research*, 36(5–7), 543–562. <https://doi.org/10.1177/0278364917707024>.
- Wolpaw, J. R., Ramoser, H., McFarland, D. J., & Pfurtscheller, G. (1998). EEG-based communication: improved accuracy by response verification. *IEEE Transactions on Rehabilitation Engineering*, 6(3), 326–333. <https://doi.org/10.1109/86.712231>.
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791. [https://doi.org/10.1016/S1388-2457\(02\)00057-3](https://doi.org/10.1016/S1388-2457(02)00057-3).
- Yin, Y. H., Fan, Y. J., & Xu, L. D. (2012). EMG and EPP-integrated human-machine interface between the paralyzed and rehabilitation exoskeleton. *IEEE Transactions on Information Technology in Biomedicine*, 16(4), 542–549. <https://doi.org/10.1109/TITB.2011.2178034>.
- Ying, R., Weisz, J., & Allen, PK. (2018) Grasping with your brain: A brain-computerinterface for fast grasp selection. In *Robotics Research*, SpringerInternational Publishing, pp. 325–340. https://doi.org/10.1007/978-3-319-51532-8_20.
- Zajac, F. E. (1989) Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in Biomedical Engineering*, 17(4), 359–411. <http://e.guigon.free.fr/rsc/article/Zajac89.pdf>.
- Zhang, H., Chavarriaga, R., Khalilardali, Z., Gheorghe, L., Iturrate, I., & Millán, J. (2015). EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering*, 12(6), 066028. <https://doi.org/10.1088/1741-2560/12/6/066028>.



Joseph DelPreto is a PhD candidate in the Distributed Robotics Lab at MIT, where he works on developing systems for natural human-robot collaboration. His focus is on using biosignals and wearable sensors to improve cognitive and physical communication channels between people and robots, enabling fluid interaction and effective robotic assistance. He obtained a Master of Science degree from MIT in Electrical Engineering and Computer Science, and a Bachelor of Science degree from Columbia University in Electrical Engineering with minors in Mechanical Engineering and Psychology.



Andres F. Salazar-Gomez is a research scientist at MIT Open Learning working on neuroscience and education. Prior, he was a postdoctoral associate at the MIT Jameel World Education Lab (JWEL) working on data science education, and a postdoctoral fellow at the Distributed Robotics Lab at MIT where he explored the use of brain activity for volitional control of robots during human-robot collaborative tasks. He was a visiting student at the Miller Lab at the Picower Institute for Learning and Memory at MIT where he worked on non-human primate electrophysiology, decoder error detection, and brain-machine interface (BMI) control. The focus of his research is developing BMI for augmentative and alternative communication and human-robot interaction, and exploring ways to improve higher education in Latin America. He received his B.Sc. in Biomedical Engineering from EIA University, Colombia, and earned his Ph.D. in Computational Neuroscience from Boston University.



Stephanie Gil is currently an Assistant Professor in the School of Engineering and Applied Sciences at Harvard University, where she directs the Robotics Embedded Autonomy and Communication Theory (REACT) Lab. She was previously an Assistant Professor in the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University. Prior, she was a research scientist in the Computer Science and Artificial Intelligence Lab (CSAIL) at MIT where she also completed her Ph.D. work (2014) on multi-robot coordination and control with an emphasis on communication and intelligent information exchange. She completed her M.S. work (2009) at MIT on system identification and model learning and her B.S. at Cornell University in 2006.



Ramin M. Hasani is a PhD research assistant in Computer Science at the Institute of Computer Engineering, of TU Wien, Austria. His primary research focus is on the development of interpretable machine learning algorithms for dynamical systems. Ramin has had his brain-inspired machine learning research featured at the TEDxVienna 2018 forum. He was a visiting research scholar at Massachusetts Institute of Technology (MIT), CSAIL, investigating interpretable neural network

topologies for autonomous robotic systems. Ramin actively collaborates with the Distributed robotics lab at CSAIL of MIT, Institute of Science and technology (IST) Austria, Infineon Technologies, and IMP Austria. Ramin has completed an M.Sc. in Electronic Engineering at Politecnico di Milano (2015) and has got his B.Sc. in Electrical Engineering–Electronics at Ferdowsi University of Mashhad (2012).



Frank H. Guenther is a professor in the Departments of Speech, Language & Hearing Sciences and Biomedical Engineering at Boston University. He is also a Research Affiliate at the Picower Institute for Learning and Memory at MIT, a Faculty Member in the Harvard/MIT Speech and Hearing Bioscience and Technology Program, and a Visiting Scientist in the Department of Radiology at Massachusetts General Hospital. He is a computational and cognitive neuroscientist. His research

combines theoretical modeling with behavioral and neuroimaging experiments to characterize the neural computations underlying speech and sensorimotor control. Dr. Guenther also develops brain-computer interfaces to restore synthetic speech and other communication skills to paralyzed individuals. He earned his PhD in Cognitive and Neural Systems from Boston University.



Daniela Rus is the Andrew (1956) and Erna Viterbi Professor of Electrical Engineering and Computer Science and Director of the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT. Her research interests are in robotics and artificial intelligence. The key focus of her research is to develop the science and engineering of autonomy. She is a Class of 2002 MacArthur Fellow, a fellow of ACM, AAAI and IEEE, and a member of the National Academy of Engineering and of the American Academy of Arts and Sciences. She is the recipient of the Engelberger Award for robotics. She earned her PhD in Computer Science from Cornell University.

can Academy of Arts and Sciences. She is the recipient of the Engelberger Award for robotics. She earned her PhD in Computer Science from Cornell University.